RECOMBINANT POLYPEPTIDES FOR DIAGNOSING INFECTION WITH *TRYPANOSOMA CRUZI*

This application claims priority from U.S. Provisional Application No. 60/430,654, filed December 4, 2002, hereby incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

[0001] 1. Field of the invention

[0002] The present invention relates to recombinant polypeptides that are useful for diagnosing American trypanosomiasis, or Chagas disease. Chagas disease is caused by the infectious agent *Trypanosoma cruzi*. More particularly, the invention relates to specific combinations of recombinant *T. cruzi* polypeptides, synthesized using genetic engineering techniques, and to constructs and processes for producing the recombinant polypeptides, and to an assay and kit for detecting *T. cruzi* infection which employs the recombinant polypeptides.

[0003] 2. Background

[0004] Chagas disease is a zoonosis caused by the protozoan parasite, *Trypanosoma cruzi*. This organism is primarily transmitted through contact with its triatomine insect vectors, but transmission by transfusion of contaminated blood and congenital transmission also are important. Historically Chagas disease has been a public health problem in all of Latin America, with the exception of the Caribbean nations. The World Health Organization estimates that 16-18 million persons are chronically infected with *T. cruzi*, and that 45,000 deaths occur each year due to the illness. Infection with *T. cruzi* is life-long and specific drug treatment lacks efficacy and often causes serious side effects. Ten to thirty percent of *T. cruzi*-infected persons develop chronic symptomatic Chagas disease, and the burden of disability and mortality in the endemic countries is enormous.

[0005] An estimated 80,000 to 100,000 *T. cruzi*-infected persons now live in the United States. These immigrants pose a risk for transfusion-associated transmission of the

parasite here and in other countries to which Latin Americans have emigrated. Eight such cases have been reported in the United States, Canada, and Europe, all of which occurred in immunosuppressed patients in whom acute *T. cruzi* infection was diagnosed because of the fulminant course of the illness. Most transfusions are given to immunocompetent patients in whom acute Chagas disease would be a mild illness, and thus it is reasonable to assume that many other undetected instances of transfusion-associated transmission of *T. cruzi* have occurred in the United States and other industrialized nations. The question of whether blood donated in the United States should be screened serologically for antibodies to *T. cruzi* has been considered for at least a decade by both public and private entities involved in blood banking. A panel of experts convened in early 2000 by the American Red Cross to consider this issue recommended unanimously that our blood supply be screened serologically. Implementation of such a recommendation, however, is not an option currently because no test for *T. cruzi* infection has been cleared by the FDA for screening donated blood.

[0006] Diagnosis of *T. cruzi* infection presents problems. Demographic and clinical data are suggestive at best. Parasitologic tests, e.g., xenodianosis, hemoculture and PCR are insensitive. Other serologic tests are generally insensitive and lack specificity, as false positive reactions often occur with specimens from patients having infectious diseases, such as leishmaniasis, syphilis, or malaria; autoimmune diseases; and other parasitic and non-parasitic illinesses.

[0007] Such conventional tests include indirect immunofluorescence (IIF), indirect hemagglutination (IHA), and complement fixation (CF) tests, as well as enzyme-linked immunosorbent assays (ELISA or EIA). Due to the lack of sensitivity and specify of the three commonly used assays, when a sample has a positive result from any, the blood must be discarded. Table I shows that in a major Brazilian blood bank (Hemocentro, São Paulo, Brazil), up to 3.43% of blood donations fall into this category.

IIF	IHA	CF	% w/ Results
+	+	+	0.68%
+	-	+	0.71%
+	+	-	
-	+	+	
+	-	-	2.04%
-	+	-	
-	-	+	
		TOTAL:	3.43%

TABLE I

[0008] Commercially available ELISAs include lysate-based tests such as the Chagas Enzyme Immunoassay (EIA), available from Abbott Laboratories of Abbott Park, Illinois (the subject of FDA 510(k) Premarket Notification No. K933716, herein incorporated by reference in its entirety); the Chagas' IgG ELISA, available from Meridian Bioscience, Inc. of Cincinnati, Ohio, and its predecessor, Gull Laboratories (the subject of FDA 510(k) Premarket Notification No. K911233, herein incorporated by reference in its entirety); and the Chagas' kit (EIA method), available from Hemagen Diagnostics, Inc., of Waltham, Massachusetts (the subject of FDA 510(k) Premarket Notification No. K930272, herein incorporated by reference in its entirely). However, because these tests have less than optimal sensitivities and specificities, their use for screening donated blood would fail to detect some *T. cruzi*-infected units and also would cause substantial numbers of otherwise usable units to be discarded needlessly.

[0009] One of the present inventors has previously developed a radioimmune precipitation assay (RIPA), described in Kirchhoff LV, Gam AA, Gusmao RD, Goldsmith RS, Rezende JM, Rassi A. "Increased specificity of serodiagnosis of Chagas' disease by detection of antibody to the 72 and 90 kDa glycoproteins of Trypanosoma cruzi." J Infect Dis 1987;155:561-564, herein incorporated by reference in its entirety. This test is considered the benchmark against which other tests are measured, and it is the only current option for confirmatory testing in the United States. Unfortunately, the RIPA costs $175 per assay, and at that price, screening the approximately 13 million units of blood donated each year would cost over $2 billion.

[0010] Therefore, the present inventors have further developed recombinant assays for detection of *T. cruzi* infection. A typical recombinant polypeptide and method for assaying is described by them in U.S. Patent No. 5,876,734, No. 6,228,601, and PCT Publication No. WO 95/25797, each of which is herein incorporated by reference in its entirety. Such assays for *T. cruzi* infection based on recombinant antigens, in contrast to those utilizing native antigens (e.g., the conventional lysate-based assays), as discussed above, will be more accurate, i.e., the sensitivity and specificity will be higher.

[0011] Furthermore, the recombinant assays of the invention present manufacturing advantages over the materials for the RIPA and conventional tests. Once the molecular biology has been completed, the recombinant antigens are produced in *Escherichia coli*, thus eliminating completely any biohazard associated with growing the parasites in liquid culture. This is a substantive advantage, as many cases of laboratory-acquired *T. cruzi* infection have been reported. Additionally, recombinant antigens produced in *E. coli* are much easier to purify, quantitate, and standardize than antigen lysates produced in liquid cultures of parasites, thus facilitating the manufacture of a consistent product and simplifying compliance with governmental regulations. A final advantage lies in the fact that several of the recombinant proteins presented in this application are comprised of two to four distinct protein segments derived from separate *T. cruzi* genes. This use of hybrid recombinant proteins also facilitates manufacture of an assay in that several antigenically distinct proteins are obtained in a single purification, quantitation, and standardization run.

SUMMARY OF THE INVENTION

[0012] The present invention utilizes recombinant proteins for detecting *T. cruzi* infected blood. The invention utilizes specific polypeptide sequences that correspond to fusion proteins FP3, FP4, FP5, FP6, FP7, FP8, FP9 and FP10 as described below. Isolated polynucleotides that encode the inventive polypeptides according to the present invention are also utilized, as are cells transformed with a recombinant plasmid that expresses a polypeptide according to the invention. The present invention is similar to that which is described in U.S. Patent No. 5,876,734, herein incorporated by reference in its entirety. However, the present invention replaces the proteins in the process with the recombinant

proteins of this invention to achieve similar or superior results.

[0013] The present invention also provides a method for detecting the presence of antibodies to *T. cruzi* in an individual, comprising the steps of contacting a putative anti-*T. cruzi* antibody-containing sample from an individual with a polypeptide according to the invention that is typically attached or conjugated to a carrier molecule or attached or conjugated to a solid phase; allowing anti-*T. cruzi* and other antibodies in said sample to bind to said polypeptide; washing away unbound anti-*T. cruzi* antibodies; and adding a compound that enables detection of the anti-*T. cruzi* antibodies which are specifically bound to the polypeptide. The compound that enables detection of the anti-*T. cruzi* antibodies may be selected from the group consisting of a colorometric agent, a fluorescent agent, a chemiluminescent agent and a radionucleotide.

[0014] Also provided in accordance with the present invention is a kit for diagnosing the presence of anti-*T. cruzi* antibodies in a sample, comprising a container in which a polypeptide according to the invention is attached or conjugated to a carrier molecule or attached or conjugated to a solid phase; and directions for carrying out the method according to the invention. The kit additionally may comprise a container of a compound that binds to anti-*T. cruzi* antibodies and that renders said antibodies detectable.

[0015] Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] Fig. 1 is a description of the prior art.

[0017] Fig. 1a-1h are schematic representations of the recombinant proteins utilized in

the invention.

[0018] Fig. 2 is a bar graph showing reactivity of various blood specimens with recombinant proteins used alone or in combination as target antigens in ELISAs.

DETAILED DESCRIPTION OF THE INVENTION

[0019] Figures 1a-1h represent the recombinant proteins of the invention, with the various letters indicating known protein sequences, as follows. The Figs. are schematic diagrams of the recombinant *T. cruzi* proteins, comprised of segments A through L. Solid segments (A, C, D, F, H, I, and K) represent nonrepetitive proteins having amino acid sequences that are unrelated to each other. Saw-tooth segments (B, E, G, J, and L) represent repetitive proteins having amino acid sequences that are unrelated to each other and unrelated to those of the nonrepetitive proteins. The relative sizes and numbers of repeats in the repetitive proteins are roughly represented in the Figs. The sizes and shapes of the nonrepetitive segments bear no relation to the actual proteins.

[0020] The following information refers to Figures 1 and 1a-1h in which the recombinant proteins Ag15, FP3, FP4, FP5, FP6, FP7, FP8, FP9 and FP10 are depicted schematically. These proteins are derived from *T. cruzi*, the protozoan parasite that causes Chagas disease, and are formed from of proteins A through L as indicated, and defined herein. There are no substantive amino acid similarities among proteins A through L. Similarly there are no substantive DNA sequence similarities among the segments that encode proteins A through L. The *T. cruzi* DNA sequences that encode proteins A through L were cloned in combination into pGEX and pET plasmid vectors, such as pET-32a. Strains of *Escherichia coli* were transfected with the recombinant vectors bearing the *T. cruzi* DNA sequences, and the bacteria were incubated in liquid culture under conditions favoring synthesis of the recombinant proteins. The latter proteins were subsequently affinity-purified and then used as target antigens in ELISAs. ELISAs in which proteins Ag15, FP3, FP4, FP5, FP6, FP7, FP8, FP9, and FP10, alone or in combination are employed as target antigens are useful as sensitive and specific detectors of anti-*T. cruzi*

antibodies in blood specimens obtained from persons who are chronically infected with this parasite. The detection of such antibodies is the primary means of identifying persons who are chronically infected with *T. cruzi*.

[0021] The following paragraphs contain information relating to the naming, localization, and function of proteins A through L, as well as the corresponding GenBank accession numbers of the sequences to which they are related and relevant publications.

[0022] It should be noted that the *T. cruzi* gene segments that encode protein segments A through L generally are shortened versions of the native coding regions. In this context, the constructs that encode single segments (i.e., FP5 and FP9), as well as all the others that encode more than one segment, are all unique, because, even if the individual components from which the various recombinant proteins of this invention are known, the segments of the invention have not been combined previously as described herein.

[0023] Protein AB. This hybrid recombinant protein, also designated Ag15 [SEQ ID NO. 2] in Figure 1, is derived from the TCR27 gene of *T. cruzi* [SEQ ID NO. 1]. Protein A is the amino terminal nonrepetitive portion of the TCR27 protein, and Protein B is comprised of approximately 18 of the 14 amino acid repeats that make up the central portion of the TCR27 protein. The two native TCR27 genes sequenced contained approximately 69 and 105 of the 14-amino acid repeats.

[0024] Nucleotide sequence data that include the Ag15 DNA sequence were deposited with GenBank and EMBL databases by Keiko Otsu, John E. Donelson, and Louis V. Kirchhoff with the accession number L04603 and are described in U.S. Patent No. 5,876,734 and No. 6,228,601, issued to Louis V. Kirchhoff and Keiko Otsu (each of which is herein incorporated by reference in its entirety). These references also present DNA and inferred protein sequences that include the Ag15 DNA and inferred protein sequences. The Ag15 DNA and inferred protein sequences are additionally presented in Otsu K, Donelson JE, Kirchhoff LV. "Interruption of a Trypanosoma cruzi gene encoding a protein containing 14-amino acid repeats by targeted insertion of the neomycin phosphotransferase gene." Mol Biochem Parasitol 1993;57:317-330, herein incorporated by reference in its entirety.

[0025] Protein C. This is a calcium binding protein of *T. cruzi*, initially called 1F8 and later designated the flagellar calcium binding protein (FCaBP) [SEQ ID NO 4]. The accession number of the original 1F8 DNA sequence [SEQ ID NO 3] deposited in GenBank is K03278. The Protein C DNA and inferred protein sequences are presented in Gonzalez A, Lerner TJ, Huecas M, Sosa-Pineda B, Nogueira N, Lizardi PM. "Apparent generation of a segmented mRNA from two separate tandem gene families in Trypanosoma cruzi." Nucleic Acids Res 1985;13(16):5789-804, herein incorporated by reference in its entirety.

[0026] Fig. 1a shows a first protein (FP3) [SEQ ID NO. 22] in accordance with the invention. Specifically, FP3 corresponds essentially to the combination of Ag15 (Fig. 1), and by Protein C. The DNA sequence encoding FP3 [SEQ ID NO 21], also essentially corresponds to the sequences coding for Ag15 and Protein C.

[0027] Protein D. This is the protein core of a surface glycoprotein of *T. cruzi* that is referred to as GP72 [SEQ ID NO 6]. The accession number of the original gp72 DNA sequence [SEQ ID NO 5] deposited in GenBank is M65021. The Protein D DNA and inferred protein sequences are presented in Cooper R, Inverso JA, Espinosa M, Nogueira N, Cross GA. "Characterization of a candidate gene for GP72, an insect stage-specific antigen of Trypanosoma cruzi." Mol Biochem Parasitol 1991;49(1):45-59, herein incorporated by reference in its entirety.

[0028] Fig. 1b shows a second protein (FP4) [SEQ ID NO 8] in accordance with the invention. The DNA sequence [SEQ ID NO 7] that encodes Protein DABC which is a single continuous coding region, essentially corresponds to the DNA sequences from which it was constructed.

[0029] Protein E. This is a segment of the flagellar repetitive protein (FRA) [SEQ ID NO 10] of *T. cruzi* comprised of approximately nine repeats consisting of 68 amino acids each, shown as Fig. 1c (FP5). The accession number of the original Protein E DNA sequence [SEQ ID NO 9] deposited in GenBank is J04015. The Protein E DNA and inferred protein sequences are presented in Lafaille JJ, Linss J, Krieger MA, Souto-Padron T, de Souza W, Goldenberg S. "Structure and expression of two Trypanosoma

cruzi genes encoding antigenic proteins bearing repetitive epitopes." Mol Biochem Parasitol 1989; 35(2):127-136, herein incorporated by reference in its entirety.

[0030] Protein FGH. This is a protein [SEQ ID NO 12] encoded by a modified version of the *T. cruzi* TCR39 gene that was artificially constructed [SEQ ID NO 11], shown as Fig. 1e (FP7). The modification entailed reducing the length of the central portion of the TCR39 gene that encodes the 12-amino acid repeats. Protein F is the amino terminal nonrepetitive segment of the TCR39 protein. Protein G is comprised of approximately 13 of the 12-amino acid repeats that make up the central portion of the TCR39 protein. Protein H is the carboxy terminal nonrepetitive segment of the TCR39 protein. The accession number of the original, i.e., the unmodified, Protein FGH DNA sequence deposited in GenBank is U15616. The TCR39 DNA and inferred protein sequences, which include the entire Protein FGH sequences, are presented in Gruber A, Zingales B. "Trypanosoma cruzi: characterization of two recombinant antigens with potential application in the diagnosis of Chagas' disease." Exp Parasitol 1993;76(1):1-12, herein incorporated by reference in its entirety.

[0031] Fig. 1d shows another hybrid recombinant protein (FP6, Protein FGHE) [SEQ ID NO 14] in accordance with the invention. The DNA sequence that encodes Protein FGHE [SEQ ID NO 13], which is a single continuous coding region, essentially corresponds to the DNA sequences from which it was constructed.

[0032] Protein IJK. This is a protein [SEQ ID NO 16] encoded by a modified version of the *T. cruzi* shed acute phase antigen (SAPA) gene that was artificially constructed [SEQ ID NO 15], as shown in Fig. 1f (FP8). The modification entailed reducing the length of the central portion of the SAPA gene that consists of 12-amino acid repeats. Protein I is the amino terminal nonrepetitive segment of the SAPA protein. Protein J is comprised of approximately nine of the 12-amino acid repeats that make up the central portion of the SAPA protein. Protein K is the carboxy terminal nonrepetitive segment of the SAPA protein. The accession number of the original, i.e., the unmodified, Protein IJK DNA sequence deposited in Gen Bank is J03985. The SAPA DNA and protein sequences, which include the entire Protein IJK sequences, are presented in Affranchino JL,

Pollevick GD, Frasch ACC. "The expression of the major shed *Trypanosoma cruzi* antigen results from the developmentally-regulated transcription of a small gene family." FEBS Lett 1991;280:316-320, herein incorporated by reference in its entirety.

[0033] Protein L. This is a microtubule-associated repetitive protein (MAP) [SEQ ID NO 18] of *T. cruzi* that is comprised of approximately five repeats consisting of 38 amino acids each, as depicted in Fig. 1g (FP9). The accession number of the original Protein L DNA sequence [SEQ ID NO 17] deposited in GenBank is S68286. The Protein L DNA and inferred protein sequences are presented in Kerner N, Liegeard P, Levin MJ, Hontebeyrie-Joskowicz M. "Trypanosoma cruzi: antibodies to a MAP-like protein in chronic Chagas' disease cross-react with mammalian cytoskeleton." Experimental Parasitology 1991;73(4):451-459, herein incorporated by reference in its entirety.

[0034] Fig. 1h shows another hybrid recombinant protein (FP10, Protein IJKL) [SEQ ID NO 20] in accordance with the invention. The DNA sequence that encodes Protein IJKL [SEQ ID NO 19], which is a single continuous coding region, essentially corresponds to the DNA sequences from which it was constructed.

[0035] Additionally, combinations of the various recombinant proteins depicted in the Figs. may be used. While it is possible to combine one or more of the recombinant proteins to form longer recombinant proteins, typically more than one recombinant protein is used simultaneously. For example, simultaneous uses of FP4 and FP5, FP5 and FP6, as well as FP4 and FP6, and combinations using more than two recombinant proteins (e.g., FP4, FP6 and FP10) are considered within the scope of the present invention. It is believed that the sensitivity and specificity of the assays according to the invention are sufficient to meet FDA standards for screening the blood supply of the United States.

[0036] Additionally, as described in U.S. Patent No. 6,228,601 (herein incorporated by reference in its entirety), polypeptides need not correspond exactly over their entire lengths to be considered within the scope of the invention. For example, a wide variety of polypeptides which contain at least one epitope embodied in the polypeptides of the invention can be used in accordance with the present invention. Based on the nucleotide

sequences, polypeptide molecules also can be produced (1) that include sequence variations, relative to the naturally-occurring sequences, (2) that have one or more amino acids truncated from the naturally-occurring sequences and variations thereof, or (3) that contain the naturally-occurring sequences and variations thereof as part of a longer sequence.

[0037] In this description, polypeptide molecules in categories (1), (2) and (3) are said to "correspond" to the amino acid sequences of the recombinant proteins of the invention. Such polypeptides also are referred to as "variants." The category of variants within the present invention includes, for example, fragments and muteins of proteins A though L, as well as larger molecules that consist essentially at least one protein sequence A through L, alone or in combination with other proteins A to L.

[0038] In this regard, a molecule that "consists essentially of" protein A to L, alone or in combination with any other proteins A to L, is one that is immunoreactive with samples from persons infected with *T. cruzi*, but that does not react with samples from patients with leishmaniasis, schistosomiasis, and other parasitic and infectious diseases, with samples from patients with autoimmune disorders and other illnesses, and with specimens from normal persons.

[0039] A "mutein" is a polypeptide that is homologous to the protein to which it corresponds, and that retains the basic functional attribute--the ability to react selectively with samples from persons infected with *T. cruzi*--of the corresponding region. For purposes of this description, "homology" between two sequences connotes a likeness short of identity indicative of a derivation of the first sequence from the second. In particular, a polypeptide is "homologous" to the corresponding protein if a comparison of amino acid sequences between the polypeptide and the corresponding region reveals an identity of greater than 40%, preferably greater than 50% and more preferably 70%. Such sequence comparisons can be performed via known algorithms, such as those described in Pearson WR, Lipman DJ. "Improved tools for biological sequence comparison." Proc Natl Acad Sci USA 1988;85(8):2444-2448, herein incorporated by reference in its entirety, which are readily implemented by computer.

[0040] A fragment of a protein of the invention is a molecule in which one or more amino acids are truncated from that protein. Muteins and fragments can be produced, in accordance with the present invention, by known de novo synthesis techniques.

[0041] Also exemplary of variants within the present invention are molecules that are longer than a protein of the invention, but that contain the region or a mutein thereof within the longer sequence. For example, a variant may include a futher fusion partner in addition to the protein of the invention. Such a fusion partner may allow easier purification of recombinantly-produced polypeptides. For example, use of a glutathione-S-transferase (26 kilodaltons, GST) fusion partner allows purification of recombinant polypeptides on glutathione agarose beads.

[0042] The portion of the sequence of a such molecule other than that portion of the sequence corresponding to the region may or may not be homologous to the sequence of a protein of the invention.

[0043] It will be appreciated that polypeptides shorter than the corresponding protein of the invention but that retain the ability to react selectively with samples from persons infected with *T. cruzi* are suitable for use in the present invention. Thus, variants may be of the same length, longer than or shorter than the protein of the invention, and also include sequences in which there are amino acid substitutions of the parent sequence. These variants must retain the ability to react selectively with samples from persons infected with *T. cruzi*.

[0044] In one embodiment, the assay of the invention uses FP4 as target antigen. Table II compares the results obtained by testing 45 pre-screened Argentinean specimens in an

		RIPA	
		+	-
FP4 ELISA	+	9	0
	-	0	36

TABLE II

FP4 ELISA with those obtained by RIPA testing.

[0045] The data in Table II show that in this group of specimens, the sensitivity and specificity of the FP4 ELISA were both 100%

[0046] Similarly, the performance of an FP4+FP6 ELISA in comparison to RIPA was

		RIPA	
		+	-
FP4+ FP6 ELISA	+	10	1
	-	0	78

TABLE III

assessed by testing 89 pre-selected Guatemalan specimens.

[0047] The data shown in Table III indicate that in this group of samples, the sensitivity of the FP4+FP6 ELISA was 100% and the specificity was 98.7%.

[0048] As shown in Fig. 2, in a FP4+FP6 ELISA, performed using standard procedures, a group of previously characterized RIPA-positive samples from several Chagas-endemic countries gave a mean reactivity (absorbance) of 2.99. Thus FP4+FP6 is the preferred embodiment among the recombinant proteins tested alone and in combination in that experiment.

[0049] It should be apparent that embodiments other than those specifically described above may come within the spirit and scope of the present invention, such as recombinant proteins comprised of different combinations and/or spatial arrangements of proteins A to L. Hence, the present invention is not limited by the above description.